

## Komparasi Hasil Deteksi Objek Algoritma YOLOv3 dan YOLOv4

**Adinugraha Dharmaputra**  
Informatics Engineering of  
Telkom University, Bandung  
adinugrahad@student.telkomun  
iversity.ac.id

**Alifio Yudhistira Aji Salis**  
Informatics Engineering of  
Telkom University, Bandung  
alifioyas@student.telkomunive  
rsity.ac.id

**Kania Ardhani Putri**  
Informatics Engineering of  
Telkom University, Bandung  
kaniaardhani@student.telkomu  
niversity.ac.id

### Abstract

Deteksi objek adalah tugas yang penting dalam komputer visi, dan pendekatan YOLOv4 telah menjadi populer dalam hal ini. Dalam penelitian ini, kami mempelajari dan membandingkan fitur dan kontribusi YOLOv4, YOLOv3, dan Darknet dalam deteksi objek. Kami menguji YOLOv4, YOLOv3, dan Darknet menggunakan data berisikan sekumpulan gambar yang kami kumpulkan.

Hasil eksperimen menunjukkan bahwa YOLOv4 memberikan akurasi yang lebih tinggi daripada YOLOv3 dalam beberapa kasus. Kami juga menemukan bahwa YOLOv4 dan YOLOv3 bersama dengan Darknet memiliki keunggulan dalam hal kinerja waktu nyata, akurasi tinggi, dan efisiensi. Dalam konteks ini, kami memperoleh pemahaman lebih lanjut tentang kemampuan deteksi objek kedua algoritma ini melalui penggunaan dataset COCO.

Kesimpulannya, YOLOv4 dan YOLOv3 dengan dukungan Darknet adalah pendekatan yang kuat dalam deteksi objek. Mereka menggabungkan keunggulan dalam kinerja waktu nyata, akurasi tinggi, dan efisiensi melalui arsitektur yang terintegrasi dan teknik fusi fitur yang canggih. Penelitian ini memberikan pemahaman yang lebih dalam tentang kemampuan deteksi objek YOLOv4

dan YOLOv3, serta perbandingannya dengan pendekatan lain dalam bidang ini.

**Kata kunci:** *deteksi objek, YOLOv4, YOLOv3, Darknet, COCO, pendekatan deteksi objek, akurasi, kinerja waktu nyata, efisiensi.*

### 1. Introduction

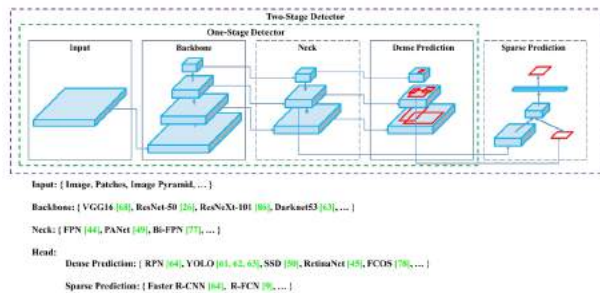
Deteksi objek adalah tugas penting dalam komputer visi yang melibatkan identifikasi dan penentuan lokasi objek dalam gambar atau video [1]. Dalam deteksi objek, metode tradisional memiliki keterbatasan dalam akurasi dan kecepatan [2]. Namun, dengan kemajuan deep learning, YOLO (You Only Look Once) menjadi pendekatan populer untuk deteksi objek [3].

YOLOv4 adalah versi terbaru dari kerangka kerja YOLO yang menggabungkan perbaikan penting untuk meningkatkan kinerja deteksi objek [1]. Implementasi YOLOv4 menggunakan kerangka kerja Darknet, yang merupakan kerangka kerja jaringan saraf sumber terbuka yang efisien [2].

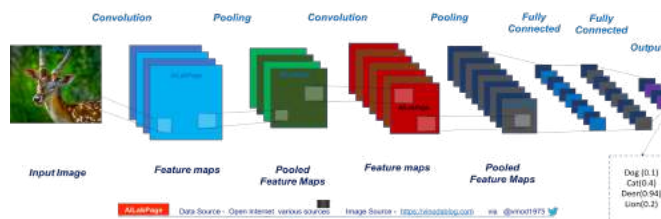
YOLOv3 (You Only Look Once v3) adalah algoritma deteksi objek real-time yang lebih akurat dan cepat dibandingkan dengan pendahulunya, YOLOv2 [4]. Dalam YOLOv3, gambar dibagi menjadi grid dan setiap sel grid memprediksi kotak pembatas objek serta probabilitas kelas [4]. Pendekatan ini memungkinkan deteksi objek dalam satu proses, cocok untuk aplikasi real-time. YOLOv3 menggunakan anchor berukuran berbeda

untuk mendeteksi objek dengan skala dan rasio aspek yang beragam [3]. Dengan fitur "skip connections" dan arsitektur "Darknet-53", YOLOv3 mencapai akurasi tinggi dan kecepatan inferensi real-time pada GPU [4]

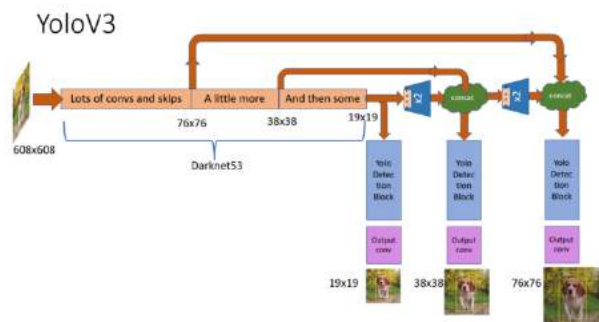
Dalam diskusi ini, kita akan mengeksplorasi fitur dan kontribusi YOLOv4, YOLOV3 dan Darknet dalam deteksi objek [1][2]. Kami akan membahas perbaikan YOLOv4 dan bagaimana Darknet memfasilitasi pengembangan dan penerapan model YOLOv4 [1][2].



Gambar 1.1 Arsitektur YOLOv4



Gambar 1.2 Arsitektur Darknet



Gambar 1.3 Arsitektur YOLOv3

## 2. Related Works

Dalam bidang deteksi objek, telah banyak karya yang diterbitkan yang berkaitan dengan proyek ini. Karya-karya ini menjelajahi pendekatan dan teknik yang berbeda untuk deteksi objek, dan memahaminya membantu untuk menyoroti kesamaan dan perbedaan dengan pendekatan yang diusulkan.

Salah satu karya terkait yang terkenal adalah Faster R-CNN (Region-based Convolutional Neural Networks) [4]. Faster R-CNN memperkenalkan konsep region proposal networks (RPNs), yang menghasilkan wilayah objek potensial untuk klasifikasi lebih lanjut. Meskipun YOLOv4 juga menggunakan pendekatan berbasis wilayah, pendekatan ini berbeda dengan arsitektur yang menyatukan deteksi dan klasifikasi dalam satu langkah, berbeda dengan pendekatan dua tahap Faster R-CNN [15].

Karya terkait lainnya adalah SSD (Single Shot MultiBox Detector) [5]. SSD memperkenalkan ide menggabungkan peta fitur dengan resolusi yang berbeda untuk mendeteksi objek pada berbagai skala dan rasio aspek. Demikian pula, YOLOv4 menggunakan jaringan piramida fitur (FPN) untuk menangkap fitur multi-skala untuk deteksi objek yang kuat. Namun, YOLOv4 meningkatkan SSD dengan menggabungkan teknik fusi fitur yang canggih dan jaringan pendukung yang lebih kuat [1].

EfficientDet adalah karya terkait lain yang menonjol [6]. Ini fokus pada pencapaian deteksi objek yang efisien dan scalable dengan menggunakan teknik scaling yang terpadu yang mengoptimalkan kedalaman dan lebar model. Meskipun YOLOv4 juga bertujuan untuk kecepatan dan akurasi yang optimal, pendekatan ini berbeda melalui pilihan arsitektur spesifik, seperti menggunakan backbone CSPDarknet53 dan menerapkan metode pemrosesan lanjutan [10].

Meskipun ada beberapa kesamaan antara YOLOv4 dan karya terkait ini, ada juga perbedaan yang jelas. Keunggulan YOLOv4 terletak pada kinerja waktu nyata dan akurasi tinggi yang dicapainya melalui arsitektur yang menyatukan,

teknik fusi fitur, dan pilihan jaringan pendukung. Penggunaan kerangka kerja Darknet juga membedakan YOLOv4, menyediakan implementasi yang efisien dan kompatibilitas dengan berbagai platform perangkat keras.

Secara keseluruhan, meskipun ada beberapa kesamaan antara YOLOv4 dan karya terkait ini, ada juga perbedaan yang jelas. Keunggulan YOLOv4 terletak pada kinerja waktu nyata dan akurasi tinggi yang dicapainya melalui arsitektur yang menyatukan, teknik fusi fitur, dan pilihan jaringan pendukung [10]. Penggunaan kerangka kerja Darknet juga membedakan YOLOv4, menyediakan implementasi yang efisien dan kompatibilitas dengan berbagai platform perangkat keras [10].

### 3. Data

Data yang digunakan dalam proyek ini adalah dataset COCO (Common Objects in Context). Jenis data ini adalah kumpulan gambar yang berisi berbagai objek yang dikategorikan ke dalam 80 kelas yang berbeda [7]. Dataset COCO ini diperoleh dari proyek Microsoft COCO yang merupakan hasil kolaborasi antara Microsoft Research dan akademisi [7]. Dataset ini terdiri dari gambar-gambar yang dikumpulkan dari berbagai sumber, termasuk penggunaan lisensi gambar dan kontribusi pengguna. Data yang digunakan dalam proyek ini mencakup sejumlah besar gambar pelatihan dan validasi dengan lebih dari satu juta objek yang dikategorikan [7].

Dalam proyek ini, kami menerapkan pendekatan menggunakan YOLOv4, YOLOv3, dan kerangka kerja Darknet untuk deteksi objek. Pada eksperimen ini kami menggunakan sekumpulan gambar untuk menguji kemampuan YOLOv3 dan YOLOv4 dalam mengenali objek. Gambar yang digunakan untuk menguji kemampuan algoritma-algoritma tersebut adalah sebagai berikut:



Gambar 1  
Dataset Jerapah dan Zebra



Gambar 2  
Dataset Anjing Manusia dan kuda

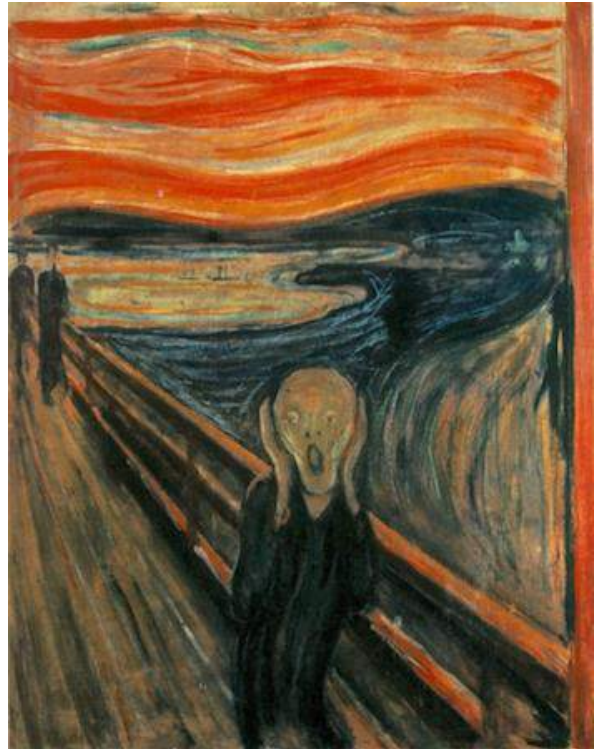




Gambar 3  
Dataset Burung



Gambar 4  
Dataset Perumahan



Gambar 5  
Dataset Gambar



Gambar 6  
Dataset Koloni Kuda

Gambar 1, 2, 3, 4, 5, dan 6 merupakan data yang digunakan untuk menguji kedua algoritma tersebut. Data yang digunakan merupakan gambar gambar yang digunakan sebagai data untuk menguji algoritma YOLOv3 dan YOLOv4.

#### 4. Metode

Pendekatan kami untuk memecahkan masalah deteksi objek melibatkan dua algoritma utama, yaitu YOLOv4 dan YOLOv3, yang diimplementasikan dalam kerangka kerja Darknet [8][9]. YOLOv3 (You Only Look Once versi 3) adalah versi sebelumnya dari YOLOv4 dan telah terbukti berhasil dalam deteksi objek real-time [1]. Kami juga mempertimbangkan YOLOv3 sebagai alternatif pendekatan.

YOLOv3 mengadopsi pendekatan yang mirip dengan YOLOv4 dalam hal pembagian gambar menjadi grid dan prediksi kotak pembatas objek serta probabilitas kelas di setiap sel grid [1]. Meskipun YOLOv3 memiliki kinerja yang sedikit lebih rendah daripada YOLOv4, pendekatannya tetap efektif dalam mendeteksi objek dalam waktu nyata dengan akurasi yang memadai [1].

Kami memilih menggunakan YOLOv4 sebagai pendekatan utama karena beberapa alasan. Pertama, YOLOv4 memiliki peningkatan signifikan dalam akurasi deteksi objek dibandingkan dengan YOLOv3 [8]. Ini penting untuk memastikan bahwa sistem deteksi objek kami memberikan hasil yang lebih akurat dan dapat diandalkan. Kedua, YOLOv4 memperkenalkan berbagai teknik baru seperti penggunaan anchor berbasis ASFF (Aggregated Feature Pyramid) dan CSPDarknet53 yang meningkatkan performa deteksi objek [8]. Dengan menggunakan YOLOv4, kami dapat memanfaatkan kemajuan terbaru dalam deteksi objek dan meningkatkan kinerja sistem kami.

Pada tahap pelatihan, kami menggunakan dataset COCO (Common Objects in Context) yang merupakan dataset populer yang berisi berbagai objek dalam berbagai konteks [7]. Kami melatih model kami dengan menggunakan metode transfer learning dengan memanfaatkan bobot yang telah

dilatih sebelumnya pada dataset COCO [8]. Dalam proses ini, kami mengoptimalkan parameter model untuk meningkatkan kinerja deteksi objek.

Pendekatan kami melibatkan langkah-langkah berikut:

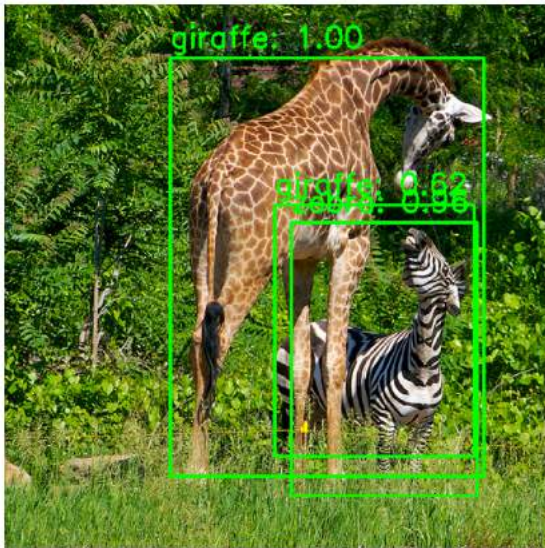
a. Pengumpulan Data: Kami menggunakan dataset COCO (Common Objects in Context) yang merupakan dataset umum yang berisi citra dengan berbagai objek dalam berbagai konteks [7].

b. Pelatihan Model: Kami menggunakan dataset COCO untuk melatih model deteksi objek kami menggunakan algoritma YOLOv4 di dalam kerangka kerja Darknet [2]. Proses pelatihan melibatkan pengoptimalan parameter model untuk mencapai kinerja deteksi objek yang maksimal.

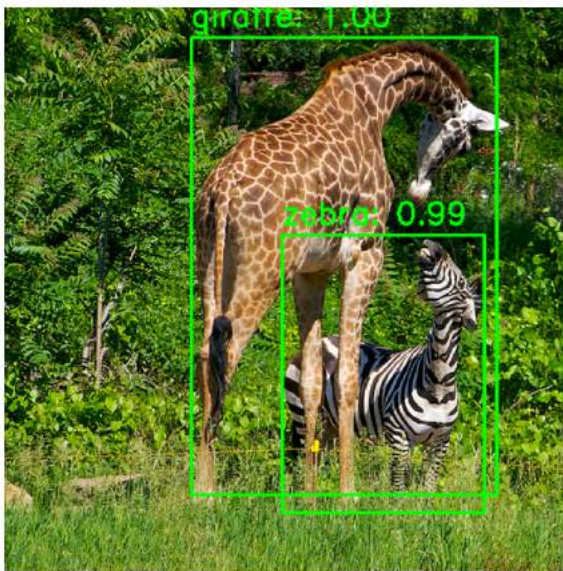
c. Komparasi Model: Setelah pelatihan selesai, kami melakukan komparasi model dan melihat bagaimana performa dari kedua algoritma tersebut dalam mendeteksi objek pada gambar yang diberikan. Kemampuan dilihat dari ketepatan algoritma dalam mendeteksi dan melabeli objek yang ada pada gambar.

#### 5. Experiments

Pada eksperimen ini, diimplementasikan algoritma YOLOv3 dan YOLOv4 untuk mendeteksi objek pada sekumpulan gambar. data yang digunakan adalah gambar 1, 2, 3, 4, 5, dan 6 yang kami kumpulkan untuk menguji kemampuan algoritma YOLOv3 dan YOLOv4 dalam mendeteksi objek pada gambar.



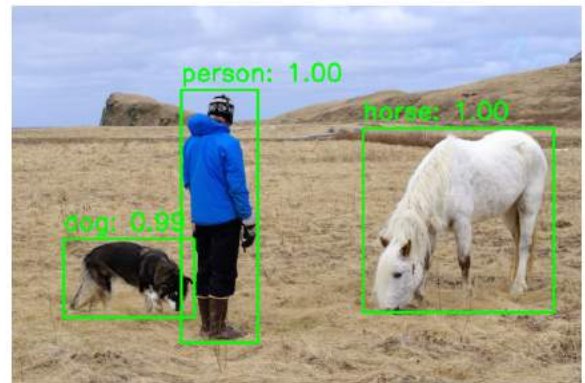
Gambar 7  
Hasil Jerapah dan Zebra Menggunakan YOLOv3



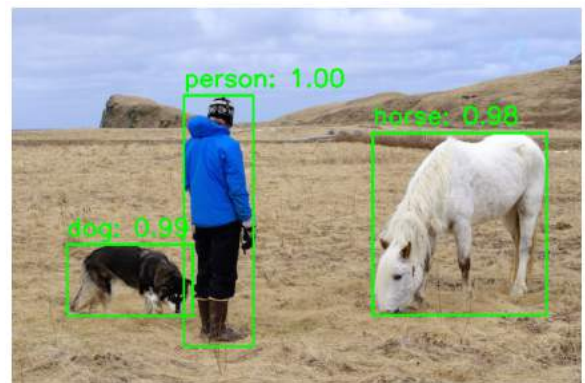
Gambar 8  
Hasil Jerapah dan Zebra menggunakan YOLOv4

Pada gambar 7, YOLOv3 mendeteksi 3 objek, 2 objek jerapah dan satu zebra. Hasil tersebut tidaklah sesuai dengan kenyataan pada gambar 1

yang hanya berisi 1 objek jerapah, dan 1 objek zebra. Sedangkan pada gambar 8, YOLOv4 hanya mendeteksi 2 objek yaitu jerapah dan zebra. YOLOv4 berhasil mendeteksi seluruh objek dengan baik.



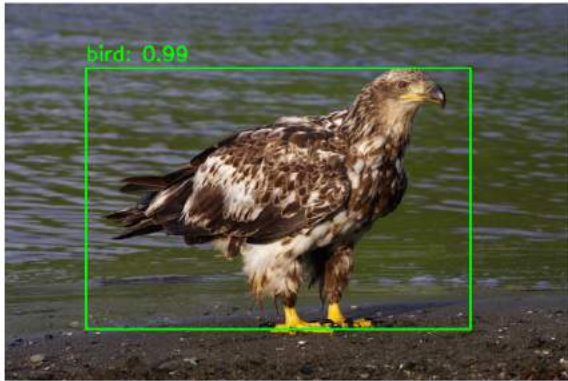
Gambar 9  
Hasil Anjing Manusia dan kuda menggunakan YOLOv3



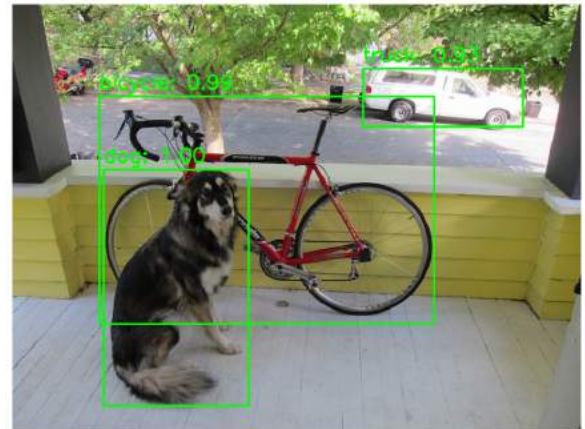
Gambar 10  
Hasil Anjing Manusia dan kuda menggunakan YOLOv4

Pada gambar 9 dan gambar 10, YOLOv3 dan YOLOv4 mendeteksi 3 objek yang terdiri dari 1 objek orang, 1 anjing dan 1 kuda. Kedua algoritma tersebut dapat mendeteksi seluruh gambar dengan benar.

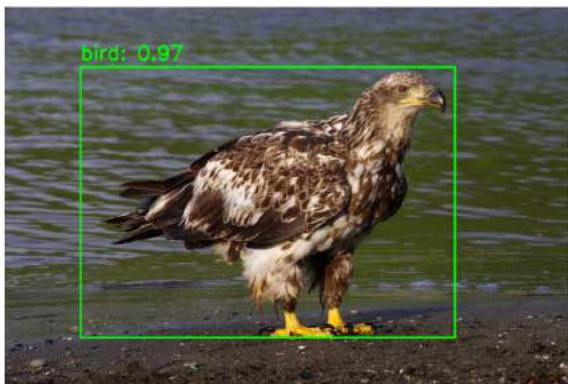




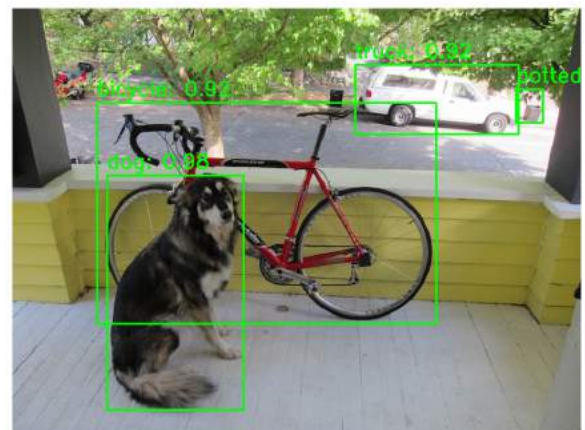
Gambar 11  
Hasil Burung menggunakan YOLOv3



Gambar 13  
Hasil Perumahan Menggunakan YOLOv3



Gambar 12  
Hasil Burung menggunakan YOLOv4



Gambar 14  
Hasil Perumahan Menggunakan YOLOv4

Pada gambar 11 dan gambar 12, YOLOv3 dan YOLOv4 mendeteksi objek burung. Kedua algoritma tersebut dapat mendeteksi seluruh gambar dengan benar.

Pada gambar 13, YOLOv3 mendeteksi 3 objek, yaitu sepeda anjing dan truk. Sedangkan pada gambar 14, YOLOv4 mendeteksi 4 objek yaitu sepeda, anjing, truk, dan tanaman pot. YOLOv4 berhasil mendeteksi lebih banyak objek, namun salah dalam melabeli tempat sampah sebagai tanaman pot.

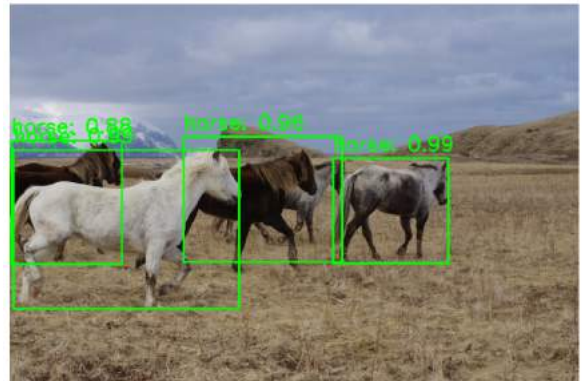


Gambar 15  
Hasil Gambar menggunakan YOLOv3

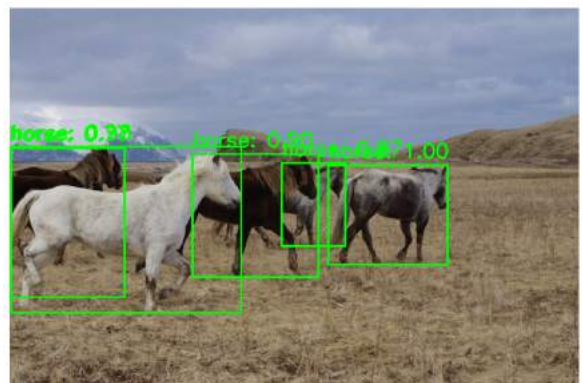


Gambar 16  
Hasil Gambar menggunakan YOLOv4  
Pada gambar 15, YOLOv3 mendeteksi kasur.

Sedangkan pada gambar 16 YOLOv4 mendeteksi koper. Namun, objek yang ada pada gambar sebenarnya adalah sebuah lukisan. YOLOv3 dan YOLOv4 tidak berhasil mendeteksi objek pada gambar dengan benar.



Gambar 17  
Hasil Koloni Kuda menggunakan YOLOv3



Gambar 18  
Hasil Koloni Kuda menggunakan YOLOv4

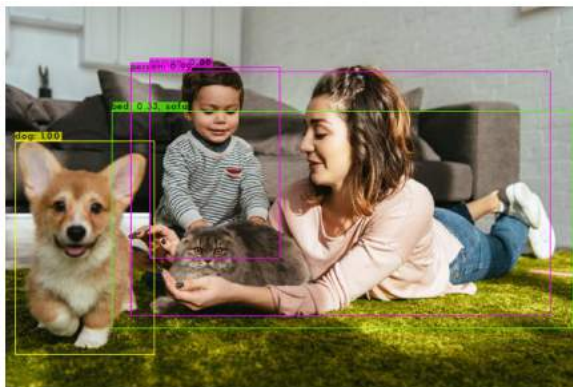
Pada gambar 17, YOLOv3 4 kuda dari 5 kuda. Sedangkan pada gambar 18 YOLOv4 mendeteksi 5 dari 5 kuda. YOLOv4 berhasil mendeteksi seluruh kuda yang ada pada gambar sedangkan YOLOv3 hanya berhasil mendeteksi 4 kuda.





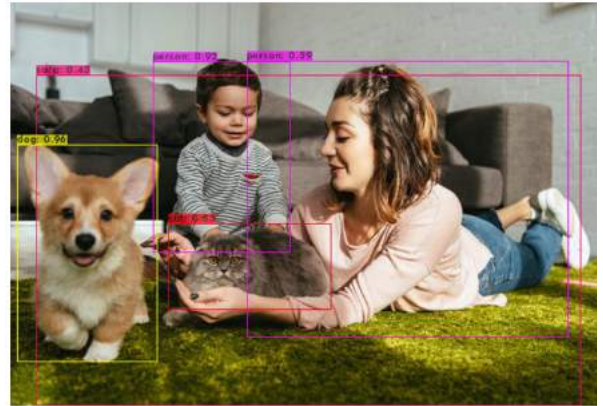
Gambar 19  
Dataset Keluarga

Gambar 19 merupakan gambar tambahan dengan banyak objek yang saya tambahkan untuk menguji kedua algoritma yang digunakan.



Gambar 20  
Hasil Keluarga Menggunakan YOLOv3

Berdasarkan gambar 20 yang merupakan hasil pengolahan gambar 19 dengan YOLOv3. Objek anjing dan 2 orang berhasil dideteksi dengan baik oleh YOLOv3, namun objek kucing dan sofa tidak berhasil dideteksi dengan baik.



Gambar 21  
Hasil Keluarga Menggunakan YOLOv4

Berdasarkan gambar 21, seluruh objek pada gambar berhasil dideteksi dengan baik oleh YOLOv4.

Berdasarkan hasil dari eksperimen tersebut YOLOv4 memiliki performa yang lebih baik dibandingkan YOLOv3 dalam mendeteksi objek. YOLOv4 dapat mendeteksi objek dengan lebih baik dan dapat melabeli objek yang ada dengan label yang lebih akurat.

## 6. Kesimpulan

Dalam proyek ini, kami menerapkan pendekatan menggunakan YOLOv3, YOLOv4 dan kerangka kerja Darknet untuk deteksi objek. Hasil dari penelitian ini dapat disimpulkan YOLOv4 menghasilkan hasil yang lebih akurat dan dapat mendeteksi objek pada gambar dengan baik. YOLOv4 menghasilkan hasil deteksi objek yang lebih minimal kesalahan dan dapat melabeli objek yang dideteksi dengan lebih akurat. Pada penelitian kedepannya diharapkan dapat dilakukan penelitian dengan dataset yang lebih besar dan analisis yang lebih andal untuk menentukan kualitas algoritma dari YOLOv3 dan YOLOv4.

## 7. Lampiran

Collabs : [TugasBesar.ipynb - Colaboratory \(google.com\)](https://colab.research.google.com/github/TugasBesar.ipynb/blob/main/Colaboratory/google.com)

Video Presentasi Google Drive :

[https://drive.google.com/file/d/182RndnmQ\\_EAsLntODnS842KE-7EqvY\\_H/view?usp=drivesdk](https://drive.google.com/file/d/182RndnmQ_EAsLntODnS842KE-7EqvY_H/view?usp=drivesdk)

## References

- [1] Bochkovskiy, A., Ryabtsev, A., & Kulachenko, A. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection.
- [2] Redmon, J. (2013). Darknet: Open Source Neural Networks in C.
- [3] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection.
- [4] Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement.
- [5] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection
- [6] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.
- [7] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single Shot MultiBox Detector.
- [8] Tan, M., Pang, R., & Le, Q. V. (2020). EfficientDet: Scalable and Efficient Object Detection.
- [9] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Girshick, R. (2014). Microsoft COCO: Common Objects in Context. European Conference on Computer Vision (ECCV).
- [10] Tan, M., & Le, Q. V. (2019). EfficientDet: Scalable and efficient object detection. Proceedings of the IEEE conference on computer vision and pattern recognition.
- [11] Redmon, J., & Farhadi, A. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934.
- [12] Redmon, J. (2013). Darknet: Open Source Neural Networks in C. <http://pjreddie.com/darknet/>
- [13] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(6), 1137-1149.
- [14] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).
- [15] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE conference on computer vision and pattern recognition
- [16] Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767.